



Station Numbers and Duration: Factors Affecting the Validity of OSCE - A Review Article

Madjid Shafiayan ¹, Azadeh Kordestani Moghaddam ^{2,*}

¹ Department of Medical Education, School of Medicine, Kurdistan University of Medical Sciences, Sanandaj, Iran

² Education Development Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

*Corresponding author: Education Development Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran. Email: kordestaniazadeh@yahoo.com

Received 2023 October 31; **Revised** 2024 June 5; **Accepted** 2024 June 10.

Abstract

Context: Objective structured clinical examination (OSCE) is one of the student assessment methods currently used in both undergraduate and postgraduate medical education. This paper reviews evidence related to key factors influenced by the number and duration of OSCE stations. The aim of our study was to broaden current knowledge of the effective components of OSCE design. Although many OSCE designers have doubts about the optimal number and duration of OSCE stations, this issue has not been adequately addressed so far.

Methods: The search was conducted through the databases of ScienceDirect, Eric, PubMed, Scopus, Ovid, and Google Scholar. English language papers, including original and review articles, without a time limit until 2022, were included. This review was conducted on papers selected using the keywords: Objective structured clinical examination, OSCE, duration of station, number of stations, utility, validity, reliability, cost, budget, finance, skill, competency.

Results: The findings have been validated by scrutinizing OSCE-related skills, reliability, validity, and cost evaluations incorporating OSCE design. These results revealed that key factors to consider in OSCE design include skills, reliability, validity, and cost.

Conclusions: Our findings could have a profound impact on the design of OSCE structures in medical education worldwide.

Keywords: OSCE Design, Station Duration, Station Number, Validity, Reliability, Cost, Clinical Skills, Competency, Assessment Methods

1. Context

Objective structured clinical examination (OSCE) is one of the most important assessment methods for both undergraduate and postgraduate students in medical education. As Zayyan stated in his 2011 paper, many authors, such as Stillman and Novack, have proposed that OSCE is a standard method to assess competencies, clinical skills, advising, and performance assessments, including data interpretation, problem-solving, training, interaction, managing unforeseen patient reactions, critical thinking, and professionalism (1). Objective structured clinical examination is recognized as a reliable and valid standardized assessment tool for assessing interpersonal communication and clinical skills (2) and has been used to teach and assess medical students in high-stakes examinations like the United States Medical Licensing

Examination (USMLE) (3). Despite the vast time and resources required to implement OSCE, it boasts fairly high validity and reliability (4, 5). Therefore, critical components such as the number and duration of test stations are important to determine in OSCE design. Harden designed the initial OSCE in 1972 in Dundee, which included eight testing stations and two rest stations. Each station was designed to assess specific competencies, such as clinical procedures, history taking, and physical examination, with 4.5 minutes allocated per station and a 30-second break between stations (6). To the best of our knowledge, various studies have concluded that OSCE design can include different numbers and durations of stations. Rushforth on behalf of Harden (1975) states that designing an OSCE typically requires 16 - 20 stations, with each station lasting 5 minutes. On the other hand, some nursing OSCE examinations have been designed with 2 to 10

stations, each ranging from 10 to 30 minutes (7). Studies regarding the number and duration of OSCE stations for medical students, residents, and fellowships have considered 6 to 42 stations, with each station lasting 4 to 20 minutes (8-13). For nursing students, at least 20 stations with durations ranging from 4 to 70 minutes have been referred to in relation to the number and duration of OSCE stations (7, 14). Although an OSCE usually includes several different stations to assess a wide range of competencies, single-station tests have also been reported (15). In Goh et al.'s 2019 study, 104 studies were reviewed, and the number of stations varied from 1 to 20. Tests with 1 to 5 stations were the most frequent (61.5%), 18.3% of studies used 6 to 10 stations, and the lowest frequency was related to tests with 11 to 20 stations (11.5%) (16). In fact, determining the number and duration of stations is required to fairly and precisely judge students' competencies. The number and duration of stations affect the test's utility, including reliability, validity, and cost (8, 17-19). To the best of our knowledge, different studies have suggested various numbers and durations of stations. Thus, the aim of our study is to determine the factors affecting the appropriate number and duration of OSCE stations.

2. Methods

The authors conducted a narrative overview literature review by examining international OSCE literature. A narrative overview was selected to achieve the aims of this study by providing an existing synthesis of knowledge related to the factors affecting the determination of the number and duration of OSCE stations. When there are one or more specific questions related to a contemporary topic and a broader view is needed, a narrative overview can be useful (20, 21).

The method described here is simple, feasible, and straightforward. The initial search was conducted through computerized databases including ScienceDirect, Eric, PubMed, Scopus, Ovid, and Google Scholar. To validate the search outcomes, we first considered specific criteria. The selected search criteria included English language papers, original and review articles, with no time limit until 2022, using the keywords: Objective structured clinical examination, OSCE, duration of station, number of stations, utility, validity, reliability, cost, budget, finance, skill, and competency. A total of 138 papers were recruited for this study.

Papers specifically referring to the quality of an OSCE were included according to the following criteria: Comparison and experience of OSCE in medical sciences related to skills, reliability, validity, and cost evaluation.

In an attempt to identify key factors, it was decided to consider inclusion criteria that included 35 studies. Figure 1 shows the process of identifying and selecting articles for the study.

3. Results

There is a vast amount of literature on the number and duration of OSCE stations. According to Khan et al. initially, a series of 18 stations, each with a duration of 5 minutes, was designed by Harden 1979 (6). The reliability, validity, clinical skills, and cost of an OSCE are influenced by the number and duration of its stations.

3.1. Clinical Skills

In 1975, Harden et al. managed the complexities of the clinical examination by incorporating various aspects of clinical scenarios, such as history taking, physical examination, and patient management problems, through 16 stations, each lasting 5 minutes (22). In 2001, Mavis designed an OSCE that included 11 stations focused on five performance skills: Interpersonal skills, physical examination skills, written records, review of systems, and problem-solving, lasting 3.5 hours to evaluate the confidence of second-year medical students (23). Townsend et al. in 2001 conducted an OSCE to identify learning deficiencies in general practice students, which included 7 stations lasting 8 - 10 minutes each, assessing skills such as history taking, physical examination, communication skills, problem-solving, information gathering, prescription writing, referral, and ethics (24). Nowadays, OSCEs typically include several stations lasting 10 to 15 minutes each to assess medical knowledge, patient care skills, and competencies like professionalism and communication (3). An OSCE was designed to evaluate pediatric students' overall perception at the end of their clerkship. This OSCE included 13 stations, each lasting 7 minutes, except for the history-taking station, which lasted 14 minutes. The skills assessed included history taking, counseling, problem-solving, diagnostic procedures, data interpretation, and photographic material analysis (12). In 2016, Gas et al. developed an OSCE for general surgery residents that included 5 stations, each lasting 15 minutes, focusing on surgical knowledge and technical skills (25). Fields et al. in 2017 designed an OSCE to compare the results of 12 years of experience with previous OSCEs in orthodontic graduates. Their OSCE comprised 13 stations, each lasting 10 minutes, and assessed five skills: Three technical, one diagnostic, and one evaluation/synthesis (26).

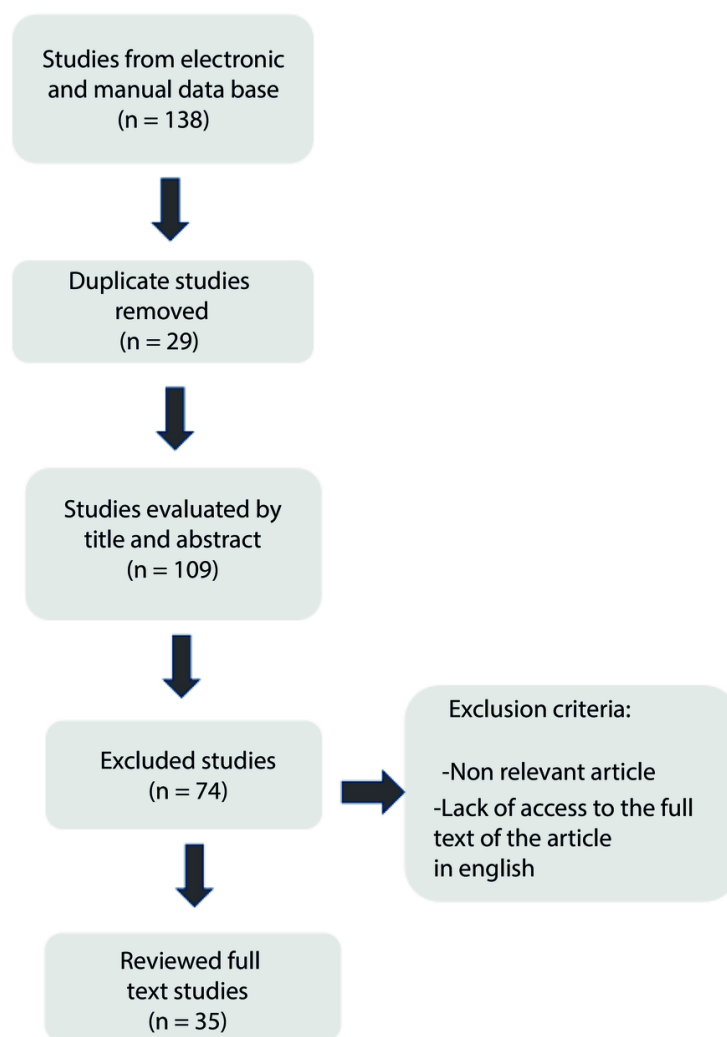


Figure 1. The flowchart of the selecting articles

3.2. Reliability

In fact, test reliability includes the stability and accuracy of students' scores with different raters over repetitions (18, 26). Although the duration of a standard OSCE station is 5 - 10 minutes, different lengths of time are recommended for a complete OSCE, such as 4 hours to obtain (≥ 0.8) reliability and 6 - 8 hours to achieve highly reliable results (26). Brannick et al. and Barman stated that if an OSCE has a large number of stations, high reliability will be achieved (8, 18). Brannick et al., in his systematic review, claims that the more OSCE

stations there are, the more reliable the results. Specifically, an OSCE with ≤ 10 stations tends to have a reliability of 0.56, while an OSCE with > 10 stations results in a reliability of 0.74. Additionally, a reliability of > 0.80 with > 25 stations has been reported in various studies (8). Khan et al. considers that 14 - 18 stations, each lasting 5 - 10 minutes, are required to obtain adequate reliability (27). Various studies consider an increased number of stations to be more effective on test reliability compared to other factors affecting test reliability, including the increased number of raters per station (27, 28).

3.3. Validity

As Messick states in his 1989 paper, validity is a key concept in the principles of theory and document estimation, which form the final judgment of the sufficiency and congruence of test scores. Ark et al. concluded that there is documented validity for only a handful of articles that establish the validity of OSCE. The theory of validity has undergone many changes, especially concerning the tools of assessment in medical education. In fact, recent theories of validity have omitted face validity (29). Additionally, evidence of reliability is considered the cornerstone of the overall validation of OSCE (30). Cohen et al. concluded that 17 of 19 stations are required to support construct validity in surgical residents (31). According to Egloff-Juras et al. and the Schoonheim study, at least 12 stations are needed to confirm the validity of the test (32). Downing wrote that to support or refute OSCE validity, sources of evidence, such as internal structure, including station analysis data, might be required (33). As Khan et al. stated, an OSCE station generally takes 5 to 10 minutes, and both the number of stations and their duration affect validity. Improving test validity in each station requires realistic and suitable time allocations for tasks (27).

3.4. Cost

Brown et al. writes that high-stakes OSCEs have higher costs compared to low-stakes OSCEs. These costs include staff, consumable equipment, travel and accommodation, venue, and patient cost divisions (19). Reznick et al. states that component costs in the four phases of the OSCE process (development, production, administration, and post-examination reporting and analysis) are similar to those in theatre production (34). Brown et al. also notes that examiner time is the main cost of an OSCE (19). Reznick et al. concludes that key factors affecting OSCE costs include the number of stations, duration, and the number of standard or simulated patients (34).

In a sequential OSCE, students participate on the first day, and those who achieve a clear pass do not participate on the second day, potentially reducing costs (19). Cookson et al. estimated that transferring an OSCE from eight long cases with 12 stations to a sequential design of four long cases with 6 stations would save approximately one-third of the original cost (35). While the initial implementation of a new OSCE is costly, subsequent administrations and shared scenarios can result in cost savings (35, 36). Additionally,

implementing an OSCE with highly trained professional actors incurs higher costs compared to using volunteers and real patients (27).

4. Discussion

The main goal of our narrative review article was to identify the effective factors influenced by the number and duration of OSCE stations. Our results indicated trends that could be useful for understanding key factors such as skills, reliability, validity, and cost, which can influence the number and duration of OSCE stations. Our results suggest that there are four elements to consider in the initial stage of OSCE design: Skills, reliability, validity, and cost. Emphasizing the importance of determining the appropriate number and duration of stations in OSCE design underscores the need for assessment methods capable of distinguishing competent medical students from incompetent ones. Barman stated that designing a reliable, valid, objective, and practicable OSCE requires great sensitivity in planning and administration, such as using a test matrix and its learning objective allocations, organizing a sufficient number of stations, training academic staff, and using proper checklists (18). Mavis designed an OSCE to examine the self-efficacy of second-year medical students with 11 stations lasting 3.5 hours, but it lacked reliability and validity (23). Numerous studies have reported various station numbers and durations to assess different skills compatibly in OSCE design (3, 12, 23, 24, 26). Pierre et al. designed an OSCE to evaluate child health medical students' perceptions at the end of their clerkship with 13 stations, each lasting 7 minutes, except for the history-taking station, which lasted 14 minutes. Therefore, considering various skills requires different station durations in OSCE design (12). Thus, it seems necessary to consider factors other than skills when determining the number of stations and test duration in OSCE development. There is controversy regarding the standard length of an OSCE station. Some researchers have obtained good reliability (≥ 0.8) with stations lasting 5 - 10 minutes each, while others recommend a 6 - 8 hour OSCE to achieve reliable results (26). Although higher reliability in OSCEs is generally achieved with more stations (18), some studies have found the opposite. Brannick et al. investigated OSCEs and found that tests with fewer than 10 stations may result in reliability greater than 0.80. Conversely, other studies showed that even with more than 25 stations, reliability of less than 0.80 was obtained (8).

Therefore, it appears that factors other than reliability affect the number and duration of stations in OSCE implementation. To remove poorly performing

OSCE stations, Auewarakul et al. analyzed the OSCE to estimate reliability and found that reliability was the basis of construct validity (30). In a study by Cohen in 1990, an OSCE was implemented for surgical residents to evaluate the construct validity of the test. Cohen et al. found that 17 of 19 stations were required to obtain construct validity, but he also considered three other elements: Reliability, skills, and cost (31). Khan et al. states that OSCE stations usually last 5 to 10 minutes each, but reliability, validity, and task allocation are essential for determining station numbers and duration (27).

Barman defines the types of validity and mentions that for an OSCE to have a high level of validity, it is important to consider the blueprint of subject matter that students are expected to achieve according to learning objectives (18). However, he does not address validity in terms of OSCE station number and duration. Direct and indirect costs are reported in a limited number of studies. OSCE costs vary due to differences in station designs across various OSCEs. In his 2015 paper, Brown et al. notes that cost resources are influenced by the time period of the study (19). Additionally, the main costs are impacted by inflation over the time period of the study.

Consequently, outdated budget estimates are not adequate resources for stakeholders. As Pell et al. state, reducing the number of stations results in decreased marginal costs, such as reductions in clinical resources, the number of simulations, support staff, and assessors, which leads to overall cost reduction (17). According to Reznick et al., multisite examinations for licensees, large-scale, high-stakes examinations require increased costs for OSCE implementation with more stations (34). Although Cookson et al. claim that if the number of stations is halved in a sequential OSCE, the cost will be reduced by one-third, it is not possible to definitively predict the amount of cost reduction because OSCE costs are influenced by the required skills, the number of assessors, standardized patients, and essential equipment (35).

Additionally, the cost of a station depends on whether real patients or professional standardized patients (SP) are used, and this cost is influenced by whether the OSCE is a new implementation or an iterated one. As mentioned previously, the focus of this study was to consider important factors related to OSCE station numbers and duration. An extensive search of databases such as PubMed, ScienceDirect, Scopus, Eric, Ovid, and Google Scholar for OSCE studies revealed four key factors.

Papers from various disciplines, including undergraduate and postgraduate medicine (e.g., internal medicine, surgery, child health, and nursing), were reviewed. It was found that factors such as skills, validity, reliability, and cost play significant roles in the initial stage of OSCE design and have a substantial influence on OSCE station numbers and duration. Similar levels of influence were observed in OSCEs across different fields like internal medicine, surgery, child health, and nursing in various universities of medical sciences. Our findings revealed that these four factors simultaneously affect OSCE station numbers and duration.

The presence of some factors with different OSCEs is due to slight differences in subjects from one discipline to another. Our reviewed studies identified four key factors compared to other studies. These studies were then categorized into four groups: Skills, reliability, validity, and cost. We selected the studies based on station numbers and duration in OSCE design developed for undergraduate and postgraduate medical sciences to scrutinize key factors. The studies elaborated in our review provide examples of factors that may be useful for developing OSCEs in the future.

Most studies, with the exception of Lind et al. in 1998, which examined the effect of length, timing, and content of the third-year surgery rotation on several clerkship and post-clerkship performance metrics, developed OSCEs to assess student competency and measure reliability and validity. However, they did not focus on the important factors influencing OSCE station numbers and duration (37). The opinions of OSCE designers offer important perspectives that can contribute to the development of more accurate, reliable, and valid tests. As mentioned in the introduction, no prior review appears to have determined the effective factors influencing OSCE station numbers and duration.

4.1. Strengths and Limitations

Investigating one of the basic needs in OSCE design, namely determining the appropriate number and duration of stations in the initial stages of test design, is one of the strengths of this study.

We are aware that our review may have two limitations. First, we could not assess all the evidence pertaining to this field. Although some progress has been made using our review, these incremental factors provide only a partial answer to OSCE development. Second, we only included studies in English, which may have limited our findings. These limitations highlight

the difficulty of collecting comprehensive data for this review.

4.2. Conclusions

The findings underscore the importance of determining the appropriate number and duration of stations in the early stages of OSCE design. The data indicate that the four variables of skills, validity, reliability, and cost have a considerable impact on OSCE design. This research suggests that investigating these factors can be instrumental in designing effective OSCE tests and help predict the utility of the designed assessments. Ultimately, this study, with its insights, can aid in developing widely used and impactful OSCE tests and pave a clearer path for future OSCE design and evaluation in the field of medical sciences. In addition to assisting in OSCE test design, the results of this study can also help predict the utility of the designed OSCEs.

4.3. Highlights

Station numbers and duration are two important components that should be considered in the early stages of designing the OSCE structure. Reliability, validity, clinical skills, and budget are crucial factors in determining the number and duration of stations. Having sufficient insight into the factors influencing the utility of the OSCE is essential for exam designers.

4.4. Lay Summary

In objective structured clinical examinations, students' clinical skills are assessed in a simulated environment. Determining the number and duration of stations are two important elements in the design of this test. The decision regarding the number and duration of stations depends on various factors. The results of this study showed that the validity of the test, the skills considered for evaluation, obtaining consistent results from repeated assessments, and the cost of the test all affect the number and duration of stations in this evaluation method.

Footnotes

Authors' Contribution: Study concept and design, A. K. M., and M. Sh.; analysis and interpretation of data, A. K. M., and M. Sh.; drafting of the manuscript, A. K. M.; critical revision of the manuscript for important intellectual content; A. K. M., and M. Sh.

Conflict of Interests Statement: There are no conflicts of interest to declare.

Data Availability: The dataset presented in the study is available on request from the corresponding author during submission or after publication.

Funding/Support: The authors received no specific funding for this work.

References

- Zayyan M. Objective structured clinical examination: The assessment of choice. *Oman Med J.* 2011;26(4):219-22. [PubMed ID: 22043423]. [PubMed Central ID: PMC3191703]. <https://doi.org/10.5001/omj.2011.55>.
- Fu CP, Chi HY, Li MW, Liu CH, Yeh JH, Wang CC, et al. Development of an objective structured clinical examination station for pediatric occupational therapy and an evaluation of its quality. *Am J Occup Ther.* 2022;76(2). [PubMed ID: 35179555]. <https://doi.org/10.5014/ajot.2022.043521>.
- Sudan R, Clark P, Henry B. Cost and logistics for implementing the american college of surgeons objective structured clinical examination. *Am J Surg.* 2015;209(1):140-4. [PubMed ID: 25454956]. <https://doi.org/10.1016/j.amjsurg.2014.10.001>.
- Moattari M, Abdollah-Zargar S, Mousavinasab M, Zare N, Beygi-Marvdast P. [Reliability and validity of OSCE in evaluating clinical skills of nursing students]. *Pejouhesh dar Pezeshki (Res Med).* 2007;31(1):55-9. Persian.
- Cullen L, Fraser D, Symonds I. Strategies for interprofessional education: The Interprofessional Team Objective Structured Clinical Examination for midwifery and medical students. *Nurse Educ Today.* 2003;23(6):427-33. [PubMed ID: 12900191]. [https://doi.org/10.1016/S0260-6917\(03\)00049-2](https://doi.org/10.1016/S0260-6917(03)00049-2).
- Khan KZ, Ramachandran S, Gaunt K, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: An historical and theoretical perspective. *Med Teach.* 2013;35(9):e1437-46. [PubMed ID: 23968323]. <https://doi.org/10.3109/0142159X.2013.818634>.
- Rushforth HE. Objective structured clinical examination (OSCE): Review of literature and implications for nursing education. *Nurse Educ Today.* 2007;27(5):481-90. [PubMed ID: 17070622]. <https://doi.org/10.1016/j.nedt.2006.08.009>.
- Brannick MT, Erol-Korkmaz HT, Prewett M. A systematic review of the reliability of objective structured clinical examination scores. *Med Educ.* 2011;45(12):1181-9. [PubMed ID: 21988659]. <https://doi.org/10.1111/j.1365-2923.2011.04075.x>.
- Cusimano MD, Cohen R, Tucker W, Murnaghan J, Kodama R, Reznick R. A comparative analysis of the costs of administration of an OSCE (objective structured clinical examination). *Acad Med.* 1994;69(7):571-6. [PubMed ID: 8018269]. <https://doi.org/10.1097/00001888-199407000-00014>.
- Sharma MK, Chandra PS, Chaturvedi SK. Team OSCE: A teaching modality for promotion of multidisciplinary work in mental health settings. *Indian J Psychol Med.* 2015;37(3):327-9. [PubMed ID: 26664082]. [PubMed Central ID: PMC4649818]. <https://doi.org/10.4103/0253-7176.162954>.
- Doig CJ, Harasym PH, Fick GH, Baumber JS. The effects of examiner background, station organization, and time of exam on OSCE scores assessing undergraduate medical students' physical examination skills. *Acad Med.* 2000;75(10 Suppl):S96-8. [PubMed ID: 11031187]. <https://doi.org/10.1097/00001888-200010001-00031>.
- Pierre RB, Wierenga A, Barton M, Branday JM, Christie CD. Student evaluation of an OSCE in paediatrics at the University of the West Indies, Jamaica. *BMC Med Educ.* 2004;4:22. eng. [PubMed ID: 15488152]. [PubMed Central ID: PMC526209]. <https://doi.org/10.1186/1472-6920-4-22>.

13. Zartman RR, McWhorter AG, Seale NS, Boone WJ. Using OSCE-based evaluation: Curricular impact over time. *J Dent Educ.* 2002;**66**(12):1323-30. [PubMed ID: [12521058](#)].
14. Walsh M, Bailey PH, Koren I. Objective structured clinical evaluation of clinical competence: An integrative review. *J Adv Nurs.* 2009;**65**(8):1584-95. [PubMed ID: [19493134](#)]. <https://doi.org/10.1111/j.1365-2648.2009.05054.x>.
15. Yap K, Sheen J, Nedeljkovic M, Milne L, Lawrence K, Hay M. Assessing clinical competencies using the objective structured clinical examination (OSCE) in psychology training. *Clinical Psychologist.* 2021;**25**(3):260-70. <https://doi.org/10.1080/13284207.2021.1932452>.
16. Goh HS, Zhang H, Lee CN, Wu XV, Wang W. Value of nursing objective structured clinical examinations: A scoping review. *Nurse Educ.* 2019;**44**(5):E1-6. [PubMed ID: [30371544](#)]. <https://doi.org/10.1097/NNE.0000000000000620>.
17. Pell G, Fuller R, Homer M, Roberts T. Advancing the objective structured clinical examination: Sequential testing in theory and practice. *Med Educ.* 2013;**47**(6):569-77. [PubMed ID: [23662874](#)]. <https://doi.org/10.1111/medu.12136>.
18. Barman A. Critiques on the objective structured clinical examination. *Ann Acad Med Singap.* 2005;**34**(8):478-82. eng. [PubMed ID: [16205824](#)].
19. Brown C, Ross S, Cleland J, Walsh K. Money makes the (medical assessment) world go round: The cost of components of a summative final year objective structured clinical examination (OSCE). *Med Teach.* 2015;**37**(7):653-9. [PubMed ID: [25923233](#)]. <https://doi.org/10.3109/0142159X.2015.1033389>.
20. Ferrari R. Writing narrative style literature reviews. *Medical writing.* 2015;**24**(4):230-5. <https://doi.org/10.1179/2047480615Z.000000000329>.
21. Green BN, Johnson CD, Adams A. Writing narrative literature reviews for peer-reviewed journals: Secrets of the trade. *J Chiropr Med.* 2006;**5**(3):101-17. [PubMed ID: [19674681](#)]. [PubMed Central ID: [PMC2647067](#)]. [https://doi.org/10.1016/S0899-3467\(07\)60142-6](https://doi.org/10.1016/S0899-3467(07)60142-6).
22. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *Br Med J.* 1975;**1**(5955):447-51. [PubMed ID: [1115966](#)]. [PubMed Central ID: [PMC1672423](#)]. <https://doi.org/10.1136/bmj.1.5955.447>.
23. Mavis B. Self-efficacy and OSCE performance among second year medical students. *Adv Health Sci Educ Theory Pract.* 2001;**6**(2):93-102. [PubMed ID: [11435761](#)]. <https://doi.org/10.1023/a:1011404132508>.
24. Townsend AH, McLlvenny S, Miller CJ, Dunn EV. The use of an objective structured clinical examination (OSCE) for formative and summative assessment in a general practice clinical attachment and its relationship to final medical school examination performance. *Med Educ.* 2001;**35**(9):841-6. [PubMed ID: [1155221](#)]. <https://doi.org/10.1046/j.1365-2923.2001.00957.x>.
25. Gas BL, Buckarma EH, Mohan M, Pandian TK, Farley DR. Objective assessment of general surgery residents followed by remediation. *J Surg Educ.* 2016;**73**(6):e71-6. [PubMed ID: [27476792](#)]. <https://doi.org/10.1016/j.jsurg.2016.07.002>.
26. Fields HW, Kim DG, Jeon M, Firestone AR, Sun Z, Shanker S, et al. Evaluation of objective structured clinical examination for advanced orthodontic education 12 years after introduction. *Am J Orthod Dentofacial Orthop.* 2017;**151**(5):840-50. [PubMed ID: [28457261](#)]. <https://doi.org/10.1016/j.ajodo.2016.10.031>.
27. Khan KZ, Gaunt K, Ramachandran S, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part II: Organisation & administration. *Med Teach.* 2013;**35**(9):e1447-63. [PubMed ID: [23968324](#)]. <https://doi.org/10.3109/0142159X.2013.818635>.
28. Govaerts MJ, van der Vleuten CP, Schuwirth LW. Optimising the reproducibility of a performance-based assessment test in midwifery education. *Adv Health Sci Educ Theory Pract.* 2002;**7**(2):133-45. [PubMed ID: [12075145](#)]. <https://doi.org/10.1023/a:1015720302925>.
29. Ark TK, Ark N, Zumbo BD. Validation Practices of the Objective Structured Clinical Examination (OSCE). In: Zumbo B, Chan E, editors. *Validity and Validation in Social, Behavioral, and Health Sciences. Social Indicators Research Series, Vol. 54.* Cham: Springer; 2014. p. 267-88. https://doi.org/10.1007/978-3-319-07794-9_16.
30. Auewarakul C, Downing SM, Praditsuwan R, Jaturatamrong U. Item analysis to improve reliability for an internal medicine undergraduate OSCE. *Adv Health Sci Educ Theory Pract.* 2005;**10**(2):105-13. [PubMed ID: [16078095](#)]. <https://doi.org/10.1007/s10459-005-2315-3>.
31. Cohen R, Reznick RK, Taylor BR, Provan J, Rothman A. Reliability and validity of the objective structured clinical examination in assessing surgical residents. *Am J Surg.* 1990;**160**(3):302-5. [PubMed ID: [2393060](#)]. [https://doi.org/10.1016/s0002-9610\(06\)80029-2](https://doi.org/10.1016/s0002-9610(06)80029-2).
32. Eglloff-Juras C, Hirtz P, Luc A, Vaillant-Corroy AS. An objective structured clinical examination (OSCE) for French dental students: Feedback after 2 years. *Dent J (Basel).* 2021;**9**(11). [PubMed ID: [34821600](#)]. [PubMed Central ID: [PMC8622547](#)]. <https://doi.org/10.3390/dj9110136>.
33. Downing SM. Validity: On meaningful interpretation of assessment data. *Med Educ.* 2003;**37**(9):830-7. [PubMed ID: [14506816](#)]. <https://doi.org/10.1046/j.1365-2923.2003.01594.x>.
34. Reznick RK, Smee S, Baumber JS, Cohen R, Rothman A, Blackmore D, et al. Guidelines for estimating the real cost of an objective structured clinical examination. *Acad Med.* 1993;**68**(7):513-7. [PubMed ID: [8323635](#)]. <https://doi.org/10.1097/00001888-199307000-00001>.
35. Cookson J, Crossley J, Fagan G, McKendree J, Mohsen A. A final clinical examination using a sequential design to improve cost-effectiveness. *Med Educ.* 2011;**45**(7):741-7. [PubMed ID: [21649707](#)]. <https://doi.org/10.1111/j.1365-2923.2010.03926.x>.
36. Walsh K, Jaye P. Cost and value in medical education. *Educ Prim Care.* 2013;**24**(6):391-3. eng. [PubMed ID: [24196592](#)]. <https://doi.org/10.1080/14739879.2013.11494206>.
37. Lind DS, Marum T, Ledbetter D, Flynn TC, Romrell LJ, Copeland E3. The effect of the duration and structure of a surgery clerkship on student performance. *J Surg Res.* 1999;**84**(1):106-11. [PubMed ID: [10334898](#)]. <https://doi.org/10.1006/jsre.1999.5624>.